

SEQUENCE LISTING

<110> ALVES, ALEXANDRA M.C.R.
RECORD, ERIC
LOMASCOLO, ANNE
SIGOILLOT, JEAN-CLAUDE
ASTHER, MARCEL
WOSTEN, HAN A.B.

<120> METHOD FOR OVERPRODUCING A SPECIFIC RECOMBINANT PROTEIN
WITH P. CINNABARINUS MONOKARYOTIC STRAINS

<130> 0508-1167

<140> 10/586,348
<141> 2006-07-14

<150> PCT/FR05/000093
<151> 2005-01-14

<150> FR 04/00366
<151> 2004-01-15

<160> 18

<170> PatentIn Ver. 3.3

<210> 1
<211> 3330
<212> DNA
<213> Pycnopus cinnabarinus

<220>
<221> CDS
<222> (128)..(310)

<220>
<221> CDS
<222> (368)..(436)

<220>
<221> CDS
<222> (490)..(610)

<220>
<221> CDS
<222> (664)..(777)

<220>
<221> CDS
<222> (833)..(896)

<220>
<221> CDS
<222> (960)..(1055)

<220>
 <221> CDS
 <222> (1114)..(1270)

<220>
 <221> CDS
 <222> (1334)..(1531)

<220>
 <221> CDS
 <222> (1592)..(1648)

<220>
 <221> CDS
 <222> (1705)..(1911)

<220>
 <221> CDS
 <222> (1968)..(2255)

<400> 1
 ctgcagacat ctggagcgcc tgtctttccc ctagtataaa tgatgtctgt ccgcaggtcc 60
 ttgaagaccg ctgcagtcgcc acttgagttt taggtaggac ctgtccacca aacctcttt 120
 tctgatac atg teg agg ttc cag tcc ctc ttc ttc ttc gtc ctc gtc tcc 169
 Met Ser Arg Phe Gln Ser Leu Phe Phe Phe Val Leu Val Ser
 1 5 10
 ctc acc gct gtg gcc aac gca gcc ata ggg cct gtg gcg gac ctg acc 217
 Leu Thr Ala Val Ala Asn Ala Ala Ile Gly Pro Val Ala Asp Leu Thr
 15 20 25 30
 ctt acc aat gcc cag gtc agc ccc gat ggc ttc gct cgc gag gcc gtc 265
 Leu Thr Asn Ala Gln Val Ser Pro Asp Gly Phe Ala Arg Glu Ala Val
 35 40 45
 gtg gtg aac ggt atc acc cct gcc cct ctc atc aca ggc aat aag 310
 Val Val Asn Gly Ile Thr Pro Ala Pro Leu Ile Thr Gly Asn Lys
 50 55 60
 gtatgtatat gctgctcgtc cctcagagct acatacatct gatccacaat cgtttag 367
 ggc gat cga ttc cag ctc aat gtc atc gac cag ttg aca aat cat acc 415
 Gly Asp Arg Phe Gln Leu Asn Val Ile Asp Gln Leu Thr Asn His Thr
 65 70 75
 atg ttg aaa aca tct agt att gtaagggttc agtttttccc gactaccatg 466
 Met Leu Lys Thr Ser Ser Ile
 80
 ttattgacca tcaccactcg tag cat tgg cac ggc ttc ttc cag caa ggc acg 519
 His Trp His Gly Phe Phe Gln Gln Gly Thr
 85 90
 aac tgg gcc gat ggt ccc gcg ttc gtg aac cag tgt ccc atc gct tcg 567
 Asn Trp Ala Asp Gly Pro Ala Phe Val Asn Gln Cys Pro Ile Ala Ser
 95 100 105 110

ggc cac tgc ttc ttg tat gac ttt caa gtt ccc gac caa gca g 610
 Gly His Ser Phe Leu Tyr Asp Phe Gln Val Pro Asp Gln Ala
 115 120

gtacgaattc cgtacacgtt tcattgcgtc gcaactaaac ctctctttac tag gg 665
 Gly
 125

act ttc tgg tac cat agc cat ctc tcc acg caa tac tgc gat ggt ttg 713
 Thr Phe Trp Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp Gly Leu
 130 135 140

agg ggg cct ttc gtc gtc tac gac ccc aac gat cct cac gct agc ctg 761
 Arg Gly Pro Phe Val Val Tyr Asp Pro Asn Asp Pro His Ala Ser Leu
 145 150 155

tat gac att gat aac g gtgagcagat catgggtatcg caatattgcg tccacttatg 817
 Tyr Asp Ile Asp Asn
 160

ctctctggca tccag ac gac act gtc att acg ctg gct gat tgg tat cac 867
 Asp Asp Thr Val Ile Thr Leu Ala Asp Trp Tyr His
 165 170

gtt gct gcc aag ctc gga cct cgc ttc cc gtacgtgtca aatgtctacg 916
 Val Ala Ala Lys Leu Gly Pro Arg Phe Pro
 175 180

agagatctca catatacgac tagactcact tcgcttgatta cag a ttt ggc tcc gat 972
 Phe Gly Ser Asp
 185

tca acc ctt atc aat gga ctt ggt cga acc act ggc ata gca ccg tcc 1020
 Ser Thr Leu Ile Asn Gly Leu Gly Arg Thr Thr Gly Ile Ala Pro Ser
 190 195 200

gac ttg gca gtt atc aag gtc acg cag ggc aag cg gtaagtatgg 1065
 Asp Leu Ala Val Ile Lys Val Thr Gln Gly Lys Arg
 205 210 215

atggatcatca ctgcacattg gctctgatac atggccttgt ttccacag c tac cgc 1120
 Tyr Arg

ttc cgc ttg gtg tgc ctt tct tgc gat ccg aac cat aca ttc agc att 1168
 Phe Arg Leu Val Ser Leu Ser Cys Asp Pro Asn His Thr Phe Ser Ile
 220 225 230

gat aat cac aca atg act ata att gag gcg gac tgc atc aac act caa 1216
 Asp Asn His Thr Met Thr Ile Ile Glu Ala Asp Ser Ile Asn Thr Gln
 235 240 245 250

ccc cta gag gtt gat tca atc cag att ttt gcc gcg cag cgc tac tcc 1264
 Pro Leu Glu Val Asp Ser Ile Gln Ile Phe Ala Ala Gln Arg Tyr Ser
 255 260 265

ttc gtg gtaggtcgta ggctcctgtc atcaagtttg cagacattct tagatacacc 1320
 Phe Val

tttttcaatg cag ctg gat gct agc cag ccg gtg gat aac tac tgg atc	1369
Leu Asp Ala Ser Gln Pro Val Asp Asn Tyr Trp Ile	
270 275 280	
cgc gca aac cct gcc ttc gga aac aca ggt ttt gct ggt gga atc aat	1417
Arg Ala Asn Pro Ala Phe Gly Asn Thr Gly Phe Ala Gly Gly Ile Asn	
285 290 295	
tct gcc atc ctg cgt tat gat ggc gca ccc gag atc gag cct acg tct	1465
Ser Ala Ile Leu Arg Tyr Asp Gly Ala Pro Glu Ile Glu Pro Thr Ser	
300 305 310	
gtc cag act act cct acg aag cct ctg aac gag gtc gac ttg cat cct	1513
Val Gln Thr Thr Pro Thr Lys Pro Leu Asn Glu Val Asp Leu His Pro	
315 320 325	
ctc tgc cct atg cct gtg gtacgtgtct caaagaacct cgatcactaa	1561
Leu Ser Pro Met Pro Val	
330	
gtgcatgtca actcatatgg tgcattgacag cct ggc agc ccc gag ccc gga ggt	1615
Pro Gly Ser Pro Glu Pro Gly Gly	
335 340	
gtc gac aag cct ctg aac ttg gtc ttc aac ttc gtgagtactg gcgcgcttcc	1668
Val Asp Lys Pro Leu Asn Leu Val Phe Asn Phe	
345 350	
gtagcacacg ttogaacaaa gcctgatacc atgcag aac ggc acc aac ttc ttc	1722
Asn Gly Thr Asn Phe Phe	
355	
atc aac gac cac acc ttt gtc ccg ccg tct gtc cca gtc ttg cta caa	1770
Ile Asn Asp His Thr Phe Val Pro Pro Ser Val Pro Val Leu Leu Gln	
360 365 370 375	
atc ctc agt ggg gcg cag gcg gct cag gac ctg gtc ccg gag ggc agc	1818
Ile Leu Ser Gly Ala Gln Ala Ala Gln Asp Leu Val Pro Glu Gly Ser	
380 385 390	
gtg ttc gtt ctt ccc agc aac tgc tcc att gag ata tcc ttc cct gcc	1866
Val Phe Val Leu Pro Ser Asn Ser Ser Ile Glu Ile Ser Phe Pro Ala	
395 400 405	
act gcc aat gcc cct gga ttc ccc cat ccg ttc cac ttg cac ggt	1911
Thr Ala Asn Ala Pro Gly Phe Pro His Pro Phe His Leu His Gly	
410 415 420	
gtacgtctgc cttccctcgc tctaaaggcg gagt cgatat ctgactccca tcacag cac	1970
His	
gcc ttc gct gtc gtc cgg agc gcc ggg agc agc gtc tac aac tac gac	2018
Ala Phe Ala Val Val Arg Ser Ala Gly Ser Ser Val Tyr Asn Tyr Asp	
425 430 435	

aac ccg atc ttc cgc gac gtc gtc agc acc ggc cag ccc ggc gac aac 2066
 Asn Pro Ile Phe Arg Asp Val Val Ser Thr Gly Gln Pro Gly Asp Asn
 440 445 450 455

gtc acg att cgc ttc gag acc aat aac cca ggc ccg tgg ttc ctc cac 2114
 Val Thr Ile Arg Phe Glu Thr Asn Asn Pro Gly Pro Trp Phe Leu His
 460 465 470

tgc cac att gac ttc cac ctc gac gca ggc ttt gct gta gtc atg gcc 2162
 Cys His Ile Asp Phe His Leu Asp Ala Gly Phe Ala Val Val Met Ala
 475 480 485

gag gac act ccg gac acc aag gcc gcg aac cct gtt cct cag gcg tgg 2210
 Glu Asp Thr Pro Asp Thr Lys Ala Ala Asn Pro Val Pro Gln Ala Trp
 490 495 500

tgc gac ttg tgc ccc atc tat gat gca ctt gac ccc agc gac ctc 2255
 Ser Asp Leu Cys Pro Ile Tyr Asp Ala Leu Asp Pro Ser Asp Leu
 505 510 515

tgagcgggat tggtactgtg acctggtgtg gggggaacat gtcgagggct ttcacgac 2315
 agggactttc aagggttgga taatatacct cacggcctgg atgactcgga cagcgtgtgg 2375
 gcgtgggtgt aactctgctt gatgttgaaa aaaggatttt atgtagaaca atttatgagc 2435
 aatcagcaat caataggatt gtgtcggttt cgacgaaatg tcttgtctcc ctgacattac 2495
 ttttgtgcga gaaatgggtc catgatacac atcat tgagc tctcaatacc aagaaggatt 2555
 acccatgtca atacccaaga tcatgtcttc gctgt ccgca atggtctcat gttgcgttga 2615
 gcagatcgca gtacgttgaa aagcgattag tatta catgc aacatgcaac atttggaagg 2675
 gggcatgcag aggttcagct cgcgtcagtc ggcca agtag cgaccttgc cgcactgcct 2735
 gttaacctga acgtatgctt cagaactccg tcggt atcga gagcgatcgt gtacgttccg 2795
 ggatagatcc attgatcccc gctctggtcg gcgctgcga tggccccgag cgtcaccggc 2855
 agcttcgcga tcgcgctttt cctaggggag aggcctgtga cccgcgtgta cgagacgagc 2915
 tgcttggtcg ggtggggcga aggccgaag gagcactca cgaagagcaa tgcgacgtaa 2975
 tccgaggtag ccttgccccg gttagtcaca cgacggaga acgtgtcgag cggcgcgagg 3035
 tcgaggaagg cggcgctctt ctgaccgcgc tgtac gaggt cggaaatcga atacgtcgat 3095
 ggcggtcctc caaagtcctg gacgttggtc gcacggccg ccgcgcctgg agctgccccaa 3155
 gaaaaatcga aggtggtgaa gtgcagtcca aagcCaaatt cgtagaccgg cgtgccggtg 3215
 taccacttgt atgtacgccc cgggttcgac gcgct tgggc gaagggtcat gtcagtcac 3275
 ggaacctgat cagcgtagat ggctgggtat tgggt gatgg gcaggcgtcc tgcag 3330

<210> 2

<211> 518

<212> PRT

<213> Pycnoporos cinnabarinus

<400> 2

Met Ser Arg Phe Gln Ser Leu Phe Phe Phe Val Leu Val Ser Leu Thr
 1 5 10 15

Ala Val Ala Asn Ala Ala Ile Gly Pro Val Ala Asp Leu Thr Leu Thr
 20 25 30

Asn Ala Gln Val Ser Pro Asp Gly Phe Ala Arg Glu Ala Val Val Val
 35 40 45

Asn Gly Ile Thr Pro Ala Pro Leu Ile Thr Gly Asn Lys Gly Asp Arg
 50 55 60

Phe Gln Leu Asn Val Ile Asp Gln Leu Thr Asn His Thr Met Leu Lys
 65 70 75 80
 Thr Ser Ser Ile His Trp His Gly Phe Phe Gln Gln Gly Thr Asn Trp
 85 90 95
 Ala Asp Gly Pro Ala Phe Val Asn Gln Cys Pro Ile Ala Ser Gly His
 100 105 110
 Ser Phe Leu Tyr Asp Phe Gln Val Pro Asp Gln Ala Gly Thr Phe Trp
 115 120 125
 Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp Gly Leu Arg Gly Pro
 130 135 140
 Phe Val Val Tyr Asp Pro Asn Asp Pro His Ala Ser Leu Tyr Asp Ile
 145 150 155 160
 Asp Asn Asp Asp Thr Val Ile Thr Leu Ala Asp Trp Tyr His Val Ala
 165 170 175
 Ala Lys Leu Gly Pro Arg Phe Pro Phe Gly Ser Asp Ser Thr Leu Ile
 180 185 190
 Asn Gly Leu Gly Arg Thr Thr Gly Ile Ala Pro Ser Asp Leu Ala Val
 195 200 205
 Ile Lys Val Thr Gln Gly Lys Arg Tyr Arg Phe Arg Leu Val Ser Leu
 210 215 220
 Ser Cys Asp Pro Asn His Thr Phe Ser Ile Asp Asn His Thr Met Thr
 225 230 235 240
 Ile Ile Glu Ala Asp Ser Ile Asn Thr Gln Pro Leu Glu Val Asp Ser
 245 250 255
 Ile Gln Ile Phe Ala Ala Gln Arg Tyr Ser Phe Val Leu Asp Ala Ser
 260 265 270
 Gln Pro Val Asp Asn Tyr Trp Ile Arg Ala Asn Pro Ala Phe Gly Asn
 275 280 285
 Thr Gly Phe Ala Gly Gly Ile Asn Ser Ala Ile Leu Arg Tyr Asp Gly
 290 295 300
 Ala Pro Glu Ile Glu Pro Thr Ser Val Gln Thr Thr Pro Thr Lys Pro
 305 310 315 320
 Leu Asn Glu Val Asp Leu His Pro Leu Ser Pro Met Pro Val Pro Gly
 325 330 335
 Ser Pro Glu Pro Gly Gly Val Asp Lys Pro Leu Asn Leu Val Phe Asn
 340 345 350
 Phe Asn Gly Thr Asn Phe Phe Ile Asn Asp His Thr Phe Val Pro Pro
 355 360 365

Ser Val Pro Val Leu Leu Gln Ile Leu Ser Gly Ala Gln Ala Ala Gln
 370 375 380
 Asp Leu Val Pro Glu Gly Ser Val Phe Val Leu Pro Ser Asn Ser Ser
 385 390 395 400
 Ile Glu Ile Ser Phe Pro Ala Thr Ala Asn Ala Pro Gly Phe Pro His
 405 410 415
 Pro Phe His Leu His Gly His Ala Phe Ala Val Val Arg Ser Ala Gly
 420 425 430
 Ser Ser Val Tyr Asn Tyr Asp Asn Pro Ile Phe Arg Asp Val Val Ser
 435 440 445
 Thr Gly Gln Pro Gly Asp Asn Val Thr Ile Arg Phe Glu Thr Asn Asn
 450 455 460
 Pro Gly Pro Trp Phe Leu His Cys His Ile Asp Phe His Leu Asp Ala
 465 470 475 480
 Gly Phe Ala Val Val Met Ala Glu Asp Thr Pro Asp Thr Lys Ala Ala
 485 490 495
 Asn Pro Val Pro Gln Ala Trp Ser Asp Leu Cys Pro Ile Tyr Asp Ala
 500 505 510
 Leu Asp Pro Ser Asp Leu
 515

<210> 3
 <211> 2527
 <212> DNA
 <213> *Pycnoporus cinnabarinus*

<400> 3
 agatctccga accagaaatg cgattgcgtt caggcccaat taagaataaa gctgcgtcag 60
 ggcagcgacg tatcttgatc catcattgac tcacgggcat cggcgtcaac accaaagcaa 120
 gctcgtccca cccataggcg tgcaccggcc ggcgtgcgcc attgaggtag atgagcgggg 180
 cgaaagtccg ccattggtag cctgtcgtg gacgcgcggc gatgaaacgt tccccaccat 240
 tgggaagaaa cgtctgcggc ccatcatccc ttcaaccgat gacaaggcgg cgtcgcgcct 300
 ttgccgcaga ggccggcggg cgacatgcac agcgaagggtc cgttgcggat gggaagcagg 360
 caatcagtgg gtgtcctacg ccgccacgat ggtcggggag cgtaggcgcc ctcccataag 420
 gcggcaagca tcatgatgct ctccgattcg ggaaacgttg tgcgatgctg gagagactct 480
 ctccgagaga ccagtgtgcg caacgttccct ggcccggaag actttaaagt gagtgtagaa 540
 gggcgagcag aggacgatca tcggattgca ggaaacatcg gcacccacag cctgggaagg 600
 atggctcttg gtagacattc gcggaagggt tcctagatgt gagcgggctt cttggatgat 660
 catgtcgtaa ctttttctga cctcgtcggg ggtaacgatg gcaggattga gcattacggg 720
 atgcctccca ttcataaacg ataaccctt ccttcagggt ggcatctccc atagagcggc 780
 acgtcttcaa ggcctagggt attcacacct ccttcgcaac atccctattc acgggtgtctg 840
 taaggaacga cttgtcatgg gatcacatga agtgacgat actgttcgcc ggtctcgcag 900
 tacagacgct agtacgggaa gtcgacatcg aagcgttcag tcaccacatg gcaaaaaagc 960
 tgcaccatac tctttatggt gagttgttcg tagttggtat acagtcattc atgagggaa 1020
 gccaccgga tagggtgtgg cgcccgcaat attcatcgcc tggcaatagt cgatgtgcgt 1080
 ccttggtcaa tgaatatcat gggtcacatg tggagacggg taaacagcgt tgactgtgaa 1140
 tccctggtgt gtgttggggc gaacagggtac gttgcaggaa caccaatatc tcttcggcag 1200

```

cccagttctt tgcgagcggc acaggcaggc atcgcgcaac agatcccagc catccggcct 1260
ctgacattcg ggatacctga agcccttcag gtacggagcg aagaggtggg ctctctgcag 1320
cgattggcgg acggatagct gtatttcctc tctcaacatt gggaagatgt gaaaggctcc 1380
atcatatagc ggctcaactc tacctcgaat gtccaaacac ggcggaataa cttattttatg 1440
tggacaaggc cgagctatga tagcttgctc ccgaagtggg taagtcccgc aatctgcggg 1500
tcaggcaaca gtctcgaaa aataagaaga atattgtagg tgcgtgtagg cgtatcgccc 1560
aaatgcgcac acacggaggc tttaggagat gaagcggccc tgagcggtaa gggagttggg 1620
tcaccgcccgc ccgacccgac tctctctctt tcccagcacc atgtctcggc gcaaacttta 1680
ccctctattg accaactcca cgagaaaagca ggaacagctt ccttgtctct catgacgtcc 1740
gcaatccaga cccttagccg gttcgttact catcgttatc cctgccgcca tggtagtggg 1800
gtcagcctgg ccagtgcgta gtcccgtctc tcttgctgca ctagagaagc cccatgagac 1860
agcgtttttt gctttatttc tgcgttttct atagacacca taggggcaaa cgatcctgca 1920
cgcccagagg tattgggctc gtcagattcc cagtttttct cctcggctctg aatcggctgc 1980
acggcagata aatcggccgg aaatgctata gcccttcata gcccgtatg agagtcgcaa 2040
aaggcttgct agtcaggctg gtcgagtggc tctcaagaag agcgtcaact tcgcgcgaca 2100
gccgcctttc agggcaagat agatcctccc atcatcccc actgcgctca gcgcgggtac 2160
cgaacaattg acttaccgac atcctccggg acgcgcaaat gctgttcgac ggaacgtaat 2220
cctcttcgct ccgcctcttt tcgctctcac gcattccgtg tgggttcgcg gacggccgct 2280
catcaggacc agaccagtct caatgtctgg tacgggcaca atgggtgacac tgcggcaact 2340
gagtaggtct ggtcaactct gtgcaccgct gcttacgctg accttcggga tactgtcctg 2400
cagacatctg gagcgccctg ctttccccta gtataaatga tgtctgtccg caggtccttg 2460
aagaccgctc gagtccact tgagttttag gtaggacctg tccaccaaac ccctctttct 2520
gatcatg

```

<210> 4
 <211> 643
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 vector

```

<400> 4
cgaccgagcg cgcgccaccc agcctatccc gcgcgggtcg ggacccaaaa taagcggggc 60
ccgcgcgcgc ccgtcggggc agcgggtgta tctacgaacg gaactgggag gcgactcgga 120
agagtttggt tagaaagggg aacaccatcg cggacggccc agtgctctgg dcagctgagc 180
gtgcattgtg ttcaattctg acctgtggca tgtaaaggaa gtgctcggga tcggaggggtg 240
gcgcgagagc ctcttcgggtg tgagattagt aactgtactg cgaagccgcg gagggggttag 300
gatgagaggt agacagggtc gcagcccagg tgcgagaagg actgcgaagg actgttcttc 360
gaccgcgcac ctgcaattgc gcgcatggat agaataagac gtcgccctcg agggggactc 420
gaccagggct ggtgggtggc cccgacggga ctggctgggc atttgcagat ggcgcgagc 480
ccaggccgcc gccgatgtgt tcatcccgtt ttgtcagtat cgatcggatc tttcgggctg 540
gggtataaaa gcgcgcgcgc cgcgctctcc ctcttctctc agcactccca tccagagcac 600
ttccctctcc catcgcatcc catcacacaa taatgcccat cac 643

```

<210> 5
 <211> 1033
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 vector


```

<400> 5
agcttctccg gccccgaatc gaacggcagg atgtgtgggc gtgtccaata ttgccatgaa 60
aatctgtcag aagtgaagccc tctcgtcacc ctgtacagct tcgctgagtt gaaaagcagg 120
gttcactctg ggctcactga tgcactgagc tcgacccggag aactaaatga ccagccggag 180
tgttactaa cttaacgccc ggtattcagg gcagcttctc tatgttgccg ctacgacgta 240
gatcaccgcc catgaacggg ggaaacgggg aggggtgcgt ttggtacgtc ttacgtctg 300
gctatgttgt attgaccagc gtctgcagaa gatgggcacg acgatgcgcc gagccggcca 360
gtgtcgtcgg atgtccactg ttgaggccat ccttttgcta gacagacgga agagctttgg 420
aggtgcgatt cctctacgaa tgggaagggg cttagatgga gagtgcacg tctgagctcc 480
ccaacacgcc ttgcccgagg gtgcgtctcc gcggacattc acctcagttc attgttctga 540
cctgcctaata tgtatagacc ggccaacaac cttgctgacg cccatcataa cagtgccttg 600
cacagagcct tcccactcag tcggcgctc cctcaatcaa tcccactaac tcgcggctc 660
tgccccttcg ccgctcgaca cgctcgttg aagagcccg gcacgggct ccgctcccc 720
cttccctccg cgctcgtcatg cagcgagcgt taatgttgct gcaggcgagc cgtaagtata 780
ttcaaaggcg tagcgaatga atagcaggcg cgcggggacc tggcacgcgc ggcatgaaca 840
tgcagacttg ggtgacgata acttgaactc agacggggcg aatgaatatc caaacgcgcg 900
ggaagaaaat aatttacggg agcctcccca ggtataaaaag cccctcaccg gctcactctt 960
tctccagtcg aacaccccag ttcaactacc cagcccttcc ttccttcgct atccttcytc 1020
acaacctgct cgc 1033

```

```

<210> 6
<211> 19
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

```

```

<400> 6
caytggcayg grttcttcc 19

```

```

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

```

```

<400> 7
gagrtggaag tcratgtgrc 20

```

```

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

```

```

<400> 8
ggataactac tggatccgcg 20

```

<210> 9
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 primer

<400> 9
 cgcagtattg cgtggagag

19

<210> 10
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 primer

<400> 10
 gacatctgga ggcctgtc

19

<210> 11
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 primer

<400> 11
 atcgaaggtt ccgatgactg acatgac

27

<210> 12
 <211> 5122
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 vector

<400> 12
 catgggatat cgcattgctg cagagctcta gact cgacgg gcccggtacc gcggccgcct 60
 taagacgcgt ggatccgcag gtgaacgcgc ctat cgggtg gatattcggg cgacggggagc 120
 ctccggaatc tgagcctcgt tactgcctag caaa ttccga atcccttcga tgatcataggg 180
 tcgaggacaa gtgatcgtct tgctacatac tcca aggtgt tgactcattc cctcgataat 240
 gaacattgtt gttgttggtt gttctctatc cgct cagtca cgcgacccca cacgtgcatg 300
 gttgaacttc gccacgcaac aaccgcatga cgacatggcg aacctaagta aaggctgagt 360
 cgtggactaa agcactccac ttacggcgga ggat gccagt ctacgtcatg aatgaagcct 420

caggtccccga agtaagggggg tacaaaagga ggggtgaaagg tggacgtttt ettaccatcc 480
 ttccacctcc cagaccacca tgccgggaat tcccagcttg ctcaaaaagg ttctgcccgt 540
 acgcccgcga aattccttcg aggtggcccc tatcgcatat atgcacgact tcaaaacatc 600
 cattctatca ttttggggtc gtacaattat tagacatgtt gtacaacgtt acattccttt 660
 cttctttttac totccggccc agtctatgta gaggtaaaagt acaagcgtcc aaaggatcag 720
 gcacttagag cgcgcgctct tgcttcgcgc cttagagcgc gccgtcctgc ttccgcgcgt 780
 agacgagcag gtccgagaca cggcgggagt agcccactc gttgtcgtac caggcaatga 840
 gcttcacgaa gctcttgctg atcgcgatgc cggggatcga tccacgcgtc ttaaggcggc 900
 cgcggtaccc cctcggaccc gtccggccgc gtccgaccgg cgggtgttgt cggcgtcggg 960
 cagtcctgct cctcggccac gaagtgcacg cagttgccgg ccgggtcgcg cagggcgaac 1020
 tccccccccc acggctgctc gccgatctcg gtcatggccg gcccgagggc gtcccggaag 1080
 ttccgtggaca cgacctccga ccactcggcg tacagctcgt ccaggccgcg caccacacc 1140
 caggccaggg tgttgctcgg caccacctgg tccctggaccg cgctgatgaa cagggtcacg 1200
 tcgtcccggg ccacaccggc tccggcgacg tccgcgcggg tgagcaccgg aacggcactg 1320
 gtcaacttgg ccattgcatg tgatggcat tatgtgtgat gggatgcgat gggagaggga 1380
 agtgctctgg atgggagtg tggagaaaga gggagacggc gggcggcgcg cctttttatc 1440
 ccacgcccga aagatccgat cgatactgac aaaaagggat gaacacatcg gcggcggcct 1500
 ggactgcgcg ccattctgca atgcccagcc agtcctcgtc ggcgccacca ccagccctgg 1560
 tcgagtcctc ctcgagggcg acgctctatt ctatccatgc gcgcaattgc aggtgcgcgc 1620
 tcgaagaaca gtccctcgca gtcctctcgc cacctgggct gcgacctgt ctacctca 1680
 tcctaaacccc tccgcggctt cgcagtacag ttactaatct cacaccgaag aggtctctgc 1740
 gccacccctc gatcccgagc acgttcctta catgcacag cgtcagaatt gaacacaatg 1800
 cagtcarat cagatccccg ggaattcgta atcatgggtc tagctgtttc ctgtgtgaaa 1860
 ttgttatccg ctcaaatc cacaacat acgagccgga agcataaagt gtaaaagcct 1920
 ggggtgcctaa tgagtgcgt aactcacatt aattgcgttg cgctcactgc ccgctttcca 1980
 gtccggaaac ctgtcgtgcc agctgcatta atgaatccgc caacgcgcgg ggagaggcgg 2040
 tttgcgtatt gggcgctctt ccgcttctc gctcactgac tgcgtgcgct cggctcgttcg 2100
 gctcggcgga gcggtatcag ctcaactcaa ggcggtaata cgggtatcca cagaatcagg 2160
 ggataacgca ggaagaaca tgtgagcaaa aggcagcaaa aaggccagga accgtaaaaa 2220
 ggccgcggtt ctggcgcttt tccataggct ccgcctccct gacgagcatc aaaaaaatcg 2280
 acgctcaagt cagaggtggc gaaaccgac aggaatataa agataccagg cgtttccccc 2340
 tggaaagctc ctcgtgcgt ctcctgttcc gaccctgcgc cttaccggat acctgtccgc 2400
 cttctccct cccggaagcg tggcgcttcc tcatagctca cgctgtaggt atctcagttc 2460
 ggtgtaggtc gttcgtcca agctgggctg tgtgcacgaa ccccccgttc agcccagacc 2520
 ctgcgcctta tccggttaact atcgtcttga gtccaaaccg gtaagacacg acttatcgcc 2580
 actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga 2640
 gttcttgaag tgggtgccta actacggcta cactagaagg acagtatttg gtatctgcgc 2700
 tctgctgaag ccagttacct tccgaaaaag agttggtagc tcttgatccg gcaaaacaa 2760
 caccgctggg agcgggtggt tttttgtttg caagcagcag attacgcgc gaaaaaagg 2820
 atctcaagaa gatcctttga tctttcttac ggggtctgac gctcagtgga acgaaaaact 2880
 acgttaaggg attttgggtc tgagattatc aaaaaggatc ttcacctaga tctttttaaa 2940
 taaaaaatga agttttaaat caatctaaag tataatagag taaacttggt ctgacagtta 3000
 ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt 3060
 tgccctgact cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag 3120
 tgctgcaatg ataccgcgag acccacgctc accggctcca gatattatcag caataaaacca 3180
 gccagccgga agggccgagc gcagaagtgg tccgcaact ttatccgcct ccattccagtc 3240
 tattaattgt tgcggggaag ctgagtaga tagtccgcca gttaatagt tgcgcaacgt 3300
 tggtgccatt gctacaggca tctgtgtgtc acgctcgtcg tttgggtatg cttcattcag 3360
 ctccggttcc caacgatcaa ggcgagttac atgaccccc atgttgtgca aaaaagcggg 3420
 tagtctcttc ggtctccga tctgtgtcag aagtgaagtt gccgcagtg tatcactcat 3480
 ggttatggca gcaactgata attctcttac tgtcatgcca tccgtaagat gcttttctgt 3540
 gactggtgag tactcaacca agtcattctg agaaatagtg atgcggcgac cgagttgtct 3600
 ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa aagtgtcat 3660
 cattggaaaa cgttcttccg ggcgaaaaat ctcaaggatc ttaccgctgt tgagatccag 3720
 ttcgatgtaa cccactcgtg caccacactg atcttcagca tcttttactt tcaccagcgt 3780
 ttctgggtga gcaaaaaacg gaaggcaaaa tgcgcgcaaaa aagggaataa gggcgacacg 3840
 gaaatgttga atactcatat tcttctcttt tcaatattat tgaagcattt atcagggtta 3900

```

ttgtctcatg agcggatata tatttgaatg tatttagaaa aataaataaa taggggttcc 3960
gogcacatct ccccgaaaag tgccacctga cgtctaagaa accattatta tcatgacatt 4020
aacctataaa aataggcgta tcacgaggcc ctttcgtctc gcgcgtttcg gtgatgacgg 4080
tgaaaacctc tgacacatgc agctcccga gacggtcaca gcttgtctgt aagcggatgc 4140
cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggctggct 4200
taactatgcg gcatcagagc agattgtact gagagtgcac catatgcggt gtgaaatacc 4260
gcacagatgc gtaaggagaa aataccgcat caggcggcat tggccattca ggctgcgcaa 4320
ctgttgggaa gggcgatcgg tgcgggcctc ttgcctatta cgccagctgg cgaaaggggg 4380
atgtgctgca aggcgattaa gttgggtaac gccagggttt tcccagtcac gacgttgtaa 4440
aacgacggcc agtgccaagc ttgcatgctt gcaggtcgac gaccgagcgc gcgccacca 4500
gcctatcccg cgcgggtcgg gacccaaaat aagcggggcc cgccgcgccc cgtcgggcga 4560
gcgggtgtat ctacgaacgg aactgggagg cgactcggaa gagtttggtt agaaagggga 4620
acaccatcgc ggacggccca gtgctctggd cagctgagcg tgcattgtgt tcaattctga 4680
cctgtggcat gtaaggaacg tgctcgggat cggagggttg cgcgagagcc tcttcggtgt 4740
gagattagta actgtactgc gaagccgcgg aggggttagg atgagaggta gacagggtcg 4800
cagcccaggc gcgagaagga ctgcgaagga ctggtcttcg accgcgcacc tgcaattcgc 4860
cgcatggata gaatagagcg tcgcccctga gggggactcg accagggtcg gtggtggcgc 4920
ccgacgggac ttgctgggca tttgcagatg gcgcgcagtc caggccgcgc ccgatgtgtt 4980
catccggttt tgtcagtatc gatcggtatc ttggggcggt ggtataaaaag cgcgcgcgcc 5040
gccgtctccc tctttctcca gcactcccat ccagaacact tccctctccc atcgcatccc 5100
atcacacaat aatgcccatac ac                                     5122

```

<210> 13

<211> 5490

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
vector

<400> 13

```

agcttctccg gccccgaatc gaacggcagg atgtgtgggc gtgtccaata ttgccatgaa 60
aatctgtcag aagtgaagcc tctcgtaacc ctgtaacagt tcgctgagtt gaaaagcagg 120
gttcatcttg ggctcactga tgcactgagc tcgacccgag aactaaatga ccagccggag 180
tgttactaa cttaacgcgc ggtattcagg gcagcttctc tatgttgccc ctacgacgta 240
gatcaccgcc catgaacggg ggaacagggg aggggtgcgt ttggtacgtc tttacgtctg 300
gctatgttgt attgaccagc gtctgcagaa gatgggcacg acgatgcgcc gagccgggca 360
gtgtcgtcgg atgtccactg ttgaggccat cttt tgcta gacagacgga agagctttgg 420
aggtgcgatt cctctacgaa tgggaagggg cttagatgga gagtgaacag tctgagctcc 480
ccaacacgcc ttgcgcgagg gtgcgtctcc gcggaattc acctcagttc attgttctga 540
cctgcctaata tgtatagacc ggccaacaac cttgctgacg cccatcataa cagtgccttg 600
cacagagcct tccactcag tcggcgccct cctcaatcaa tcccactaac tcgcccggctc 660
tgccccttcg ccgctcgaca cgtcgtttgg aagagcccgg gcacgggcgt ccgctcccc 720
cttccctccg cgtcgtcatg cagcgagcgt taatgttgct gcaggcgagc cgtaagtata 780
ttcaaaggcg tagcgaatga atagcaggcg cgcggggacc tggcacgcgc ggcataaaca 840
tgcagacttg ggtgacgata acttgaactc agacgcggcg aatgaatatc caaacgcgcg 900
ggaagaaaat aatttacggg agcctcccca ggtataaaaag cccctcaccg gctcactctt 960
tctccagtcg aacaccccag ttcaactacc cagcccttcc ttccttcgct atccttcytt 1020
acaacctgct cgccatggga tatcgcatgc ctgcagagct ctgactcga cgggcccggg 1080
accgcccggc ccttaagacg cgtggatccg caggtgaacg cgcctatcgg tgggatatte 1140
gggacgaggg agcctcggca atctgagcct cgttactgcc tagcaaatte ggaatccctt 1200
cgatgtcata gggtcgcgga caagtgatcg tcttcttaca tactccaagg tgttgactca 1260
ttccctcgat aatgaacatt gttgttgttg tttgttctct atccgctcag tcacgcgacc 1320
ccacacgtgc atggttgaaac ttgcgcacgc aacaaccgca tgacgacatg gcgaacctaa 1380
gtaagggtcg agtcgtggac taaagcaact cactttacgg cgaggatgcc agtctacgtc 1440
atgaatgaag cctcaggtcc cgaagtaagg gggtaaaaaa ggagggtgaa aggtggacgt 1500

```

tttcttacc	tccttcacc	tcccagacca	ccatgcg	cgagg	aattcccagc	ttgctcaaaa	1560
aggttctgcc	cgtacgccc	cgaaattcct	tccaggtggc	ccctatcgca	tacatgcacg		1620
acttcaaaac	atccattcta	tcattttggg	atcgta caat	tattagacat	gttggtacaac		1680
gttacattcc	tttcttcttt	tactctccgg	cccagt ctat	gtagaggtaa	agtacaagcg		1740
tccaaaggat	caggcactta	gagcgcgccc	tcttgcttcg	ccgcttagag	cgcgcgctcc		1800
tgcttcgccc	cgtagacgag	caggctcgag	acacggcggg	agtagcccca	ctcgttgtcg		1860
taccaggcaa	tgagcttcac	gaagctcttg	ctgat Cgcga	tgccggggat	cgatccacgc		1920
gtcttaaggc	ggccgcggta	ccccctcgga	cccgctcgggc	cgcgtcgga	cggcgggtgt		1980
ggtcggcgct	ggtcagtcct	gctcctcggc	cacga agtgc	acgcagttgc	cggccgggtc		2040
gcgcaggggc	aactcccgcc	cccacggctg	ctcgc Cgatc	tcgggtcatgg	ccggcccgga		2100
ggcgctcccg	aagttcgtgg	acacgacctc	cgacctctcg	gcgtacagct	cgtccaggcc		2160
gcgcacccac	acccaggcca	gggtgtgtgc	cggca Ccacc	tggtcctgga	ccgcgctgat		2220
gaacagggtc	acgtcgtccc	ggaccacacc	ggcga agtgc	tcctccacga	agtcccgga		2280
gaacccgagc	cggtcgggtc	agaactcgac	cgctc Cggcg	acgtcgcgcg	cggtgagcac		2340
cggaaaggca	ctgggtcaact	tgcccatgca	tggtgatggg	cattatgtgt	gatgggatgc		2400
gatgggagag	ggaagtgtc	tggtatgggag	tgctg gagaa	agagggagac	ggcggcgggc		2460
gcgcctttta	taccacgccc	cgaaagatcc	gatcgatact	gacaaaacgg	gatgaacaca		2520
tcggcgggcg	cctggactgc	gcgccatctg	caaat gccca	gccagtccc	tcggcgcca		2580
ccaccagccc	tggtcgagtc	ccccctcgag	gcgac gctct	attctatcca	tcgcgcgaat		2640
tgccaggtgc	cggtcgaaga	acagtccttc	gcagt Ccttc	tcgcacctgg	gctgcgaccc		2700
tgctactctc	tcctcctaac	ccctccgccc	cttcg Cagta	cagttactaa	tcctacaccc		2760
aagaggctct	cgcgccaccc	tcggatcccc	agcac gttcc	ttacatgcc	cagcgtcaga		2820
attgaacaca	atgcacgtca	ratcagatcc	ccgggaattc	gtaatcatgg	tcatagctgt		2880
ttcctgtgtg	aaattgttat	ccgctcacia	ttcca Cacia	catacgagcc	ggaagcataa		2940
agtgtaaagc	ctgggggtgc	taatgagtga	gctaa Ctcac	attaattgcg	ttgcgctcac		3000
tgcccgcctt	ccagtcggga	aacctgtcgt	gccag ctgca	ttaatgaatc	ggccaacgcg		3060
cggggagagg	cggtttgctg	attggggcgt	cttcc gcttc	ctcgctcact	gactcgctgc		3120
gctcggctct	tcggctgcgg	cgagcgggat	cagct cactc	aaaggcggt	atacggttat		3180
ccacagaatc	aggggataac	gcaggaaaga	acatg tgagc	aaaaggccag	caaaaaggcca		3240
ggaaccgtaa	aaaggccgcg	ttgctggcgt	ttttc catag	gctccgcccc	cctgacgagc		3300
atcacaaaaa	tcgacgctca	agtcagaggt	ggcga aaacc	gacaggacta	taaagatacc		3360
aggcgctttc	ccctggaagc	tcctctcgtc	gctct cctgt	tcgcacctg	ccgcttaccg		3420
gatacctgtc	cgcctttctc	ccttcgggaa	gcgtggcgct	ttctcatagc	tcacgctgta		3480
ggatctctag	tcgggtgtag	gtcgttcgct	ccaag ctggg	ctgtgtgcac	gaaccccccg		3540
ttcagcccca	cctgcgcgcc	ttatccggta	actat cgtct	tgagtcacac	ccggtaagac		3600
acgacttatc	gccactggca	gcagccactg	gtaac aggat	tagcagagcg	aggatgttag		3660
gcgggtgtac	agagtctctg	aagtgggtggc	ctaac tacgg	ctacactaga	aggacagtat		3720
ttggtatctg	cgtctctgctg	aagccagtta	ccttc Cggaaa	aagagttggt	agctcttgat		3780
ccggcaaaaa	aaccaccgct	ggtagcgggtg	gtttt tttgt	ttgcaagcag	cagattacgc		3840
gcagaaaaaa	aggatctcaa	gaagatcctt	tgat Cttttc	tacgggggtct	gacgctcagt		3900
ggaacgaaaa	ctcacgttaa	gggattttgg	tcatgagatt	atcaaaaagg	atcttcacct		3960
agatcctttt	aaattaaaaa	tgaagtttta	aatca atcta	aagtatatat	gagtaaaact		4020
ggctctgacag	ttaccaatgc	ttaatcagtg	aggca cctat	ctcagcgatc	tgtctatttc		4080
gttcatccat	agttgcctga	ctccccgtcg	tgtagataac	tacgatacgg	gagggcttac		4140
catctggccc	cagtgctgca	atgataccgc	gagac Cccacg	ctcaccggct	ccagatttat		4200
cagcaataaa	ccagccagcc	ggaagggccg	agcgcagaag	tggtcctgca	actttatccg		4260
cctccatcca	gtctattaat	tggtgcgggg	aagct agagt	aagtagttcg	ccagttaata		4320
gtttgcgcaa	cgttgctgcc	attgctacag	gcactcgtgg	gtcacgctcg	tcgtttggta		4380
tggtctcatt	cagctccggg	tcaccaacgat	caaggcgag	tacatgatcc	cccatgttgt		4440
gcaaaaaagc	ggttagctcc	ttcggtctct	cgat cgttgt	cagaagtaag	ttggccgcag		4500
tggtatcact	catgggttatg	gcagcaactgc	ataat tctct	tactgtcatg	ccatccgtaa		4560
gatgcttttc	tgtgactggg	gagtactcaa	ccaagtcatt	ctgagaatag	tgtatgcggc		4620
gaccgagttg	ctcttgcccc	gcgtcaatac	gggta aatac	cgcgccacat	agcagaactt		4680
taaaagtgtc	catcattgga	aaacgttctt	cggggcgaaa	actctcaagg	atcttaccgc		4740
tggtgagatc	cagttcgatg	taaccactc	gtgca cccaa	ctgatcttca	gcactcttta		4800
ctttcaccag	cgtttctggg	tgagcaaaaa	caggaaaggca	aaatgccgca	aaaaagggaa		4860
taagggcgac	acggaaatgt	tgaatactca	tactcttctt	ttttcaatat	tattgaagca		4920
tttatcaggg	ttattgtctc	atgagcggat	acataatttga	atgtatttag	aaaaataaac		4980

```

aaataggggt tccgcgcaca tttccccgaa aagtgcacac tgacgtctaa gaaaccatta 5040
ttatcatgac attaacctat aaaaataggg gtatcacgag gccctttcgt ctgcgcggtt 5100
tcgggtgatga cgggtgaaaac ctctgacaca tgcagctccc ggagacgggt acagcttgctc 5160
tgtaagcgga tgcggggagc agacaagccc gtcagggcgc gtcagcgggt gttggcgggt 5220
gtcggggctg gcttaactat gcggcatcag agcagattgt actgagagtg caccatatgc 5280
ggtgtgaaat accgcacaga tgcgtaagga gaaaa taccg catcaggcgc cattcgccat 5340
tcaggctgcg caactgttgg gaagggcgat cgggtcgggc ctcttcgcta ttacgccagc 5400
tggcgaaaagg gggatgtgct gcaaggcgat taagt tgggt aacgccaggg ttttcccagt 5460
cacgacgttg taaaacgacg gccagtgcc 5490

```

<210> 14

<211> 6983

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
vector

<400> 14

```

catgggatat cgcattgctg cagagctcta gagtgcacgg gcccggtacc gcggccgcct 60
taagacgcgt ggatccgcag gtgaacgcgc ctatcgggtg gatattcggg cgacggggagc 120
ctcggcaatc tgagcctcgt tactgcctag caaat tcgga atcccttcga tgtcataggg 180
tcgcggaaca gtgatcgtct tgcctacatac tccaa ggtgt tgactcattc cctcgataat 240
gaacattgtt gttgttgttt gttctctatc cgctcagtc cgcgacccca cacgtgcatg 300
gttgaacttc gccacgcaac aaccgcatga cgaca tggcg aacctaahta aaggctgagt 360
cgtggactaa agcactccac tttacggcga ggtatgccgt ctacgtcatg aatgaagcct 420
caggtcccga agtaaggggg taaaaaagga ggggtgaaagg tggacgtttt cttaccatcc 480
ttccacctcc cagaccacca tgcggggaat tccca gcttg ctcaaaaagg ttctgcccgt 540
acgcccgcga aattccttcg aggtggcccc tatcg catac atgcacgact tcaaaacatc 600
cattctatca ttttgggata gtacaattat tagacatgtt gtacaacgtt acattccttt 660
cttcttttac tctccggccc agtctatgta gaggt aaagt acaagcgtcc aaaggatcag 720
gcacttagag cgcgcctct tgettgcgcg cttagagcgc gccgtcctgc ttgcgcgctg 780
agacgagcag gtcgcagaca cggcgggagt agcccactc gttgtcgtac caggcaatga 840
gcttcacgaa gctcttgctg atcgcgatgc cggggatcga tccacgcgtc ttaaggcggc 900
cgcgggtacc cctcggacc gtccggccgc gtcgggaccg cgggtgttgg cggcgtcggg 960
cagtcctgct cctcggccac gaagtgcacg cagttgccgg ccgggtcgcg caggggcgaa 1020
tcccgcctcc acggctgctc gccgatctcg gtcatggcgg gcccgagggc gtcccgggag 1080
ttcgtggaca cgacctccga ccactcggcg tacagctcgt ccaggccgcg caccacaccc 1140
caggccaggg tgttgtccgg caccacctgg tccctggaccg cgctgatgaa cagggtcacg 1200
tcgtcccga ccacaccggc gaagtctctc tcca cgaagt cccgggagaa cccgagccgg 1260
tcgggtccaga actcgaccgc tccggcgacg tgcgcgcgg tgagcaccgg aacggcactg 1320
gtcaacttgg ccatgcatgg tgatgggcat tatgtgtgat gggatgcatg gggagaggga 1380
agtgtctctg atgggagtgc tggagaaaga gggagacggc gggcggcgcg ccttttatac 1440
ccacgcccga aagatccgat cgatactgac aaaa cgggat gaacacatcg gcggcgccct 1500
ggactgcgcg ccactctgca atgcccagcc agtc cgtcg ggcgccacca ccagccctgg 1560
tcgagtcctc ctcgaggggc acgctctatt ctatccatgc gcgcaattgc aggtgcgcgg 1620
tcgaagaaca gtctctcgca gtctctctcg cacctgggct gcgacctgt ctacctctca 1680
tcctaaccct tccgcggctt cgcagtacag ttactaatct cacaccgaag aggtctctcg 1740
gccacctctc gatcccgagc acgttcctta catgccacag cgtcagaatt gaacacaatg 1800
cacgtcarat cagatccccg ggaattcgta atcatggta tagctgtttc ctgtgtgaaa 1860
ttgttatccg ctcaacaatt cacacaacat acgagccgga agcataaagt gtaaaagcctg 1920
gggtgcctaa tgagtgcgt aactcacatt aattgcgttg cgctcactgc ccgctttcca 1980
gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg 2040
tttgcttatt gggcgtctt ccgcttctc gctcactgac tcgctgcgct cggctcgttcg 2100
gctgcggcga gcgggtatcag ctactcaaa ggcggtaata cggttatcca cagaatcagg 2160

```

ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	2220
ggcccggttg	ctggcggttt	tccataggct	ccgcccccct	gacgagcatc	acaaaaatcg	2280
acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	agataaccagg	cgtttccccc	2340
tggaagctcc	ctcgctcgct	ctcctgttcc	gacctgccc	cttaccggat	acctgtccgc	2400
ctttctccct	tggggaagcg	tggcgcttcc	tcatagctca	cgctgtaggt	atctcagttc	2460
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	ccccccgttc	agcccgaccg	2520
ctgcgcctta	tccggtaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	2580
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	2640
gttcttgaag	tgggtggccta	actacggcta	actaagaagg	acagtatttg	gtatctgcgc	2700
tctgctgaag	ccagttacct	tcggaaaaaag	agttggtagc	tcttgatccg	gcaaacacaa	2760
caccgctggg	agcgggtggg	tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	2820
atctcaagaa	gataccttga	tcttttctac	ggggctgac	gctcagtggg	acgaaaactc	2880
acgttaaggg	attttgggtc	tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	2940
ttaaaaatga	agttttaaat	caatctaaag	tatatatgag	taaacttggg	ctgacagtta	3000
ccaatgctta	atcagtggag	cacctatctc	agcga tctgt	ctatttctgt	catccatagt	3060
tgctgactc	cccgctgtgt	agataactac	gatacgggag	ggcttaccat	ctggccccag	3120
tgtgcaatg	atcccgcgag	acccacgctc	accggctcca	gatttatcag	caataaacca	3180
gccagccgga	agggccgagc	gcagaagtgg	tctgtcaact	ttatccgctc	ccatccagtc	3240
tattaattgt	tgccgggaag	ctagagtaag	tagttcgcca	gttaatagtt	tgcgcaacgt	3300
tgttgccatt	gctacaggca	tcggtggtgt	acgctcgctc	tttgggtatg	cttcattcag	3360
ctccggttcc	caacgatcaa	ggcgagttac	atgatcccc	atgttgtgca	aaaaagcggg	3420
tagctccttc	ggctcctcga	tcggtgtcag	aagtaagttg	gccgcagtgt	tatcactcat	3480
ggttatggca	gcactgcata	attctcttac	tgtcatagcc	tccgtaagat	gcttttctgt	3540
gactggtgag	tactcaacca	agtcattctg	agaatagttg	atgcggcgac	cgagttgctc	3600
ttgcccggcg	tcaatacggg	ataataccgc	gccacatagc	agaactttaa	aagtgtcat	3660
cattggaaaa	cgttcttcgg	ggcgaaaact	ctcaaaggatc	ttaccgctgt	tgagatccag	3720
ttcgatgtaa	cccactcggt	cacccaactg	atcttcagca	tcttttactt	tcaccagcgt	3780
ttctgggtga	gcaaaaaacag	gaaggcaaaa	tgccgcaaaa	aagggaataa	gggcgcacag	3840
gaaatgttga	atactcatat	tcttcttttt	tcaatattat	tgaagcattt	atcactctat	3900
ttgtctcatg	agcggatata	tatttgaatg	tatttagaaa	aataaacaaa	taggggttcc	3960
gcgcacattt	ccccgaaaag	tgccacctga	cgtct aagaa	accattatta	tcatgacatt	4020
aacctataaa	aataggcgta	tcacgaggcc	ctttcgtctc	gcgcgtttcg	gtgatgacgg	4080
tgaaaacctc	tgacacatgc	agctcccggg	gacgggtcaca	gcttgtctgt	aagcggatgc	4140
cgggagcaga	caagcccgtc	agggcgcgct	agcgggtgtt	ggcgggtgtc	ggggctgggt	4200
taactatgcy	gcacagagc	agattgtact	gagagtgac	catatgoggg	gtgaaatacc	4260
gcacagatgc	gtaaggagaa	aataccgcat	caggcgccat	tcgccattca	ggctgcgcaa	4320
ctgttgggaa	gggcgatcgg	tgccggcctc	ttcgctatta	cgccagctgg	cgaaaggggg	4380
atgtgtgca	aggcgattaa	gttgggtaac	gccagggttt	tcccagtcac	gacgttgtaa	4440
aacgacggcc	agtgccaaag	ttagatctcc	gaacagaaaa	tgcgattgcy	ttcaggccca	4500
attaagaata	aagctgcgtc	agggcagcga	cgtatcttga	tccatcattg	actcaccggc	4560
atcggcgctc	acaccaaagc	aagctcgctc	caccataggg	cgtgcacccg	ccggcgctgc	4620
ccattgaggt	acatgagcgg	ggcgaaaagc	gccattgggt	agccctgtcg	tggacgcgcg	4680
gcgatgaaac	gtttcccaac	attgggaaga	aacgtctgcy	gcccacatc	ccttcaccgg	4740
atgacaaggc	ggcgctcgcy	ctttgccgca	gaggccggcg	ggcgacatgc	acagcgaagg	4800
tccgttgccg	atgggaagca	ggcaatcagt	gggtgtccta	cgccgccacg	atggtcgggg	4860
agcgtaggcg	ccctcccata	agggcggaag	catcatgatg	ctctccgatt	cggaagcct	4920
ggtgcatgcy	tggagagact	ctctccgaga	gaccagtgtg	cgcaacgttc	ctggccttga	4980
agactttaaa	gtgagtgtag	aagggcgagc	agaggacgat	catcggattg	caggaaccat	5040
cggcatcctc	agcctgggaa	ggatgggtct	tggtagacat	tcgcggaagg	tgtcctagat	5100
gtgagcgggc	ttcttggatg	atcatgtcgt	aactttttct	gacctcgtcg	gtggtagcga	5160
tggcaggatt	gagcattacg	gtatgcctcc	cattcataaa	cgataacccc	ttccttcagg	5220
ttgggtcatct	ccatagagcg	gcacgctctc	aaggcctagg	ctattcacac	ctccttcgca	5280
acatccctat	tcacgggtgtc	tgtaaagAAC	gacttgtcat	gggatcacat	gaagtgcagc	5340
atactgttcg	ccgggtctcg	agtacagacg	ctagtacggg	aagtcgacat	ccaagcgttc	5400
agtcaccata	tggcaaaaaa	gtgcacacat	actctttatg	gtgagttgtt	cgtgagtggt	5460
atacagtcac	tcatgagggg	atgcccaccc	gatagggtgt	ggcgcccgca	atattcatcg	5520
cctggcaata	gtcgatgtgc	gtccttggtc	aatgaatatc	atgggtcaca	tgtggagacg	5580
gttaaacagc	gttgactgtg	aatccctggg	gtgtgttggg	ccgaacaggt	acgttgacgg	5640

```

aacaccaata tctcttcggc agcccagttc tttgcgagcg gcacaggcag gcatcgcgca 5700
acagatccca gccatccggc ctctgacatt cggga tacct gaagcccttc aggtacggag 5760
cgaagaggtg ggctctctgc agcgattggc ggacggatag ctgtatttcc tctctacca 5820
ttgggaagat gtgaaaggct ccatcatata gcggc tcaac tctacctga atgtccaaac 5880
acggcgggaa tacttattta tgtggacaag gccga gctat gatagcttgc tcccgaagtt 5940
ggtaagtccc gcaatctgcg gttcaggcaa cagtc tccga aaaataagaa gaatattgta 6000
ggtgcgtgta ggcgtatcgc ccaaatgcgc acaca cggag gctttaggag atgaagcgcc 6060
cgtgagcggg aagggagttg gttcacgcgc gcccc gaccg actctctctc tttcccagca 6120
tcatgtctcg gcgcaaactt taccctctat tgacc aactc cagagaaaag caggaacagc 6180
ttccttgtct ctcatgacgt ccgcaatcca gacct ttagc cggttcgtta ctcatcgta 6240
tccctgccgc catcgtagt gagtacgct ggcca gtgcg tagtcccgtc tctcttgctg 6300
cactagagaa gccccatgag acagcgtttt ttgct ttatt tctgctgttt ctatagacac 6360
cataggggca aacgatcctg cagcccaga ggtat tgggc tcgtcagatt cccagttttt 6420
ctcctcggtc tgaatcggct gcacggcaga taaat cggcc ggaaatgcta tagcccttca 6480
tagcccgcta tgagagtcgc aaaaggcttg tcagt caggt cggtcgagtg gctctcacga 6540
agagcgtcaa cttcgcgcga cagcgcctt tcagg gcaag atagatctc ccatcatccc 6600
ctactgcgtc cagcgcggg accgaacaat tgact taccg acatcctccg ggacgcgcaa 6660
atgctgttcg acggaacgta atcctcttcg tcccgcctct tttcgtctc acgcattccg 6720
tgtggttcgc gcgacggccg ctcatcagga ccaga ccagt ctcaatgtct ggtaccggca 6780
caatggtgac actgcggcaa ctgagtaggt ctggt cactc tgggtgcaccg tcgcttacgc 6840
tgaccttcgg gatactgtcc tgcagacatc tggag cgctt gtctttcccc tagtataaat 6900
gatgtctgtc cgcaggtcct tgaagaccgc tcgag tccca cttgagtttt aggtaggacc 6960
tgttccctca caaccctct ttc 6983

```

<210> 15

<211> 4395

<212> DNA

<213> *Pycnoporus cinnabarinus*

<400> 15

```

gagcttactg gatcttccag agaaatcggt ggaga ggtcg gccggtcagc ctccaccgaca 60
tttgacgtgg ccgatgatcc tgtggatgcc atcgt gttgg atcctgagct tgcattccatc 120
gccaagcgcg tcaaagctga ggtgggaagg cagggaggca caccagttcc cgaaggaggc 180
ggacctgaga ttgtgacgct caagattata tggaa acctc atccgctgaa ccccaacggc 240
cgtccggaac tctgggctat gaagcagaga cgggt aggtg aagtcgctca tcacgcctcg 300
ttcttactca ccatcttctg cagcacgaga atttc caccg gctttgttcc gaagtagcgg 360
acctcgcgag tgttcgtagt gagaacgtcg tgct t tccct cgacgggaaa cgcgtgttcc 420
cttcatctac cctcacagt gtcggtgtct gggca gaagc tgagctaggt tagtgactta 480
tcctgtgctg gacggcacga tgcttactct tcaa cagaag cttgtgacaa gatcacctac 540
caatacctgc aggaaaaataa gcgaatgcgt tccga atccg ttgctccgcc aacctatct 600
cacctcgacg acatttcccg tcagtctcca actcgcgcgc gctccccttc catcaccgag 660
ctgtccgaga atgaatccgg cgtgcagag tctggtcctg aagataaggg cacaagcact 720
ggggaggcct tcagcctgat actcgtgagc gaacggacca agggcaaggc aataaccctc 780
cgtgtgctcc caaccaccaa atgtggcgct atagt tcgca agttcctaga gaaggccggc 840
ttacaggacg aataccccga tgtcacccct gctgcgaatg ggcgaggacg cacgaagacg 900
tcagccaaga caccggcgct gacgctcgat ggagataaga tggatccgga ggcacctatt 960
ggtgatgccg acctggaaga tggggatcaa gtcgagggtg ttggtctttg atgtagcgag 1020
tgcggtggtg tacgttttct tcttgctatc ggtt tttctg ctcttgctct gttagtaagt 1080
agtataatga tggataatca cacaacaacg tatgtgttcc agggacttct ctctcagtgg 1140
gtgtgtggct gattgtacga aacatcgac aggcctttca ccgctgctcc tagcgcgaga 1200
ccacatgaac gccctcgcac gtcagtcggc ctgc cgaacg atagggcagt gccaaatgca 1260
ggcgaanaatg actcagttag gccacgcctg cgt ttaact ttagcgttct aataccctcg 1320
acctcgtag cggctcgtg agcgcggaaa gcttggcatc tacgctcttt gtccgggtccg 1380
aagccactac attgagcggc tcgccccggc ggtgacaaac tgcgaggcgc aagaatgtag 1440
ccgggcctgc ggaaaggcca tgaacaaact atgt cggccc caaccagtc taccgacacc 1500
gttctccgtg tttcagtatg ccttcagctg tcggtggggc ggggtggctcc gatattgtgta 1560
ctcggaacg ctcacagcgc tctttgtatt gccgggtatg tgaccaacgg tgcctcatg 1620

```



```

ctcttcgctt gttgatgctc cttcaggaca ccgtctggga ctctggcaag tcagctgctg 1680
ctcgccacag ttctaggaac gtctcaatgc tcctaggcgg cggttacagc aaacgccttg 1740
caccgggatg ggcctcggta cgccgcacag gcgaggctgt cctactcggc gttcgttagt 1800
agccccccat ccacgtaaga gtacctcctg cagccacccat cgtctactag cgtaccaccc 1860
acgtccactc acatcatatg ccgcccagcg cccccggact gattccgcgc tattgttgag 1920
atataagagg agtggttcgaa cggaccaagg agccataatc ccctcgagca tttcgagatc 1980
ctctccccac tgaactcctt cgccgtcacc acaaaacctc gcgtagatgt cacacttcat 2040
cgttactggg cctgtaggag gtcagactga gggcgctcct gctcccaacc gcctcgaaat 2100
caacgacttc gtcaagaatg aggagttcct ctgcgtttac gtccaggctc tcggtgcgtc 2160
gccttggcac atgtatgctc acccctatta ccatgaagct catgagccct cactacatag 2220
agatatcatg tatggactga agcaggagga actgatctcg ttcttccaga tcggtggcat 2280
tcatggattg ccatacgttg cctggagtga tgccggagcg gatgaccctg ctgagccgtc 2340
cgggtactgt acccatggct ccgtactgtt ccgcacctgg cataggcctt acgtcgcact 2400
atatgaggta agcagcttgc tagatcagac cgctaaggac gacgctgaga ctcaaaatgg 2460
ctacagcaaa tcttgacaaa gtatgctgga gagatcgctg ataagtacac ggtcgacaaa 2520
ccgcgttggc agaaggcagc ggccgacctg cgccaacctt tctgggactg ggccaagaac 2580
acgtgccttc ctactgaagt catctctctc gacaaagtca cgattacgac accagatgga 2640
cagaggacgc aagttgacaa tccactccgt cgctaaccgt tccatccgat cgaccccgac 2700
ttcccagagc catacagcaa ctggccagcg acaactgagc atccgacaag tgatggctcg 2760
gatgccaagg acaacgtgaa ggatctcact acgtaagcca attcgccata aagacgctcc 2820
tccattcatc tcaatgtata tatgtgacag tactctgaag gcggaccagc ctgatatcac 2880
gacgaagacg tataatctat tgaccagagt gcacacgtgg ccggcggttca gcaaccacac 2940
tccaggcgat ggcggcagct ccagtaacag tcttgaggcc attcacgacc acatccatga 3000
ctcagttggc ggcggaggcc agatgggaga cccgtccgtg gcaggatgt gaagtgattc 3060
ttcgcgagag acgtgactta catgtccttg taggccttga cccaatcttc ttctgcacc 3120
attgccaagt tgatcgtctt cttgcactgt ggtccgcctt gaaccccggc gtgtgggtca 3180
acagctctag ctccgaagat ggcacctaca cgatcccgcc tgactctacc gtggaccaa 3240
ctactggtgg gttcccgcac agctgtgcgc tgtggagtcg ccgttgactt ccatcactct 3300
cagcattgac gcccttcttg gatacccaaa gcacattctg gacgtccttc cagtctgctg 3360
gagtctcgcc cagccaattt ggctattctt accccgagtt taacggtctc aacctgcaag 3420
atcagaaggc tgtgaaagat cacatcgccg aggtcgtgaa cgagctctac ggtcatcgca 3480
tgccgaaaaa cttccctttc ccccgactcc aggcagtttc cgtagccaag cagggcgacg 3540
ccgtcactcc atccgtggct accgattcag tgtcgtcttc taccacacct gccgaaaatc 3600
ccgcatcccg cgaggatgcc tctgataagg acacagagcc gacgtcaat gttagggttg 3660
ccgcgccagg cgcgcacttg acctccacca agtatggga ctggactgct cgcattcatg 3720
tcaagaagta cgaagtccga ggcagcttca gcgtcctgct ctctctgggt gcaatccccg 3780
agaaccagc ggattggcgc acgagcccca actacgttgg ccgtcatcat gctttcgtga 3840
atagctcacc gcagcgctgc gctaactgcc gtggtcaagg cgacctgtc atcgaaggct 3900
tcgtccatct caacgaggcg atcgcccgcc atgcgcacct cgactccttc gatccaaccg 3960
tcgtgaggcc gtacctcagc cgcgagttgc actggggtgt gatgaaggtc agtgccata 4020
ctctgcatac gaccgtatat gtcgctaatt agatctatca aggtgaatgg caccgtcgtg 4080
ccccgcaag acgtcccgtc gtcgaggtt gtcgtcctct caactcctct tacccttctt 4140
ccgggagagc cattccctgt ccccggaacg cccgtcaatc atcatgacat caccatgga 4200
cgtcctggtg gctctacca cagcactaa gcatgctgat ggctgcccc tattgattaa 4260
acacgagtcg acctgagaac acatacaatg gatgtaatca tacttcaact ttgatgacaa 4320
tcgcttcacac attctgttcc tagcgggaca gataaccag tcaaaaaaaaa aaaaaaaaaa 4380
aaaacactgt catgc 4395

```

<210> 16

<211> 618

<212> PRT

<213> *Pycnopus cinnabarinus*

<400> 16

Met Ser His Phe Ile Val Thr Gly Pro Val Gly Gly Gln Thr Glu Gly

1

5

10

15

Ala Pro Ala Pro Asn Arg Leu Glu Ile Asn Asp Phe Val Lys Asn Glu
 20 25 30
 Glu Phe Phe Ser Leu Tyr Val Gln Ala Leu Asp Ile Met Tyr Gly Leu
 35 40 45
 Lys Gln Glu Glu Leu Ile Ser Phe Phe Gln Ile Gly Gly Ile His Gly
 50 55 60
 Leu Pro Tyr Val Ala Trp Ser Asp Ala Gly Ala Asp Asp Pro Ala Glu
 65 70 75 80
 Pro Ser Gly Tyr Cys Thr His Gly Ser Val Leu Phe Pro Thr Trp His
 85 90 95
 Arg Pro Tyr Val Ala Leu Tyr Glu Gln Ile Leu His Lys Tyr Ala Gly
 100 105 110
 Glu Ile Ala Asp Lys Tyr Thr Val Asp Lys Pro Arg Trp Gln Lys Ala
 115 120 125
 Ala Ala Asp Leu Arg Gln Pro Phe Trp Asp Trp Ala Lys Asn Thr Leu
 130 135 140
 Pro Pro Pro Glu Val Ile Ser Leu Asp Lys Val Thr Ile Thr Thr Pro
 145 150 155 160
 Asp Gly Gln Arg Thr Gln Val Asp Asn Pro Leu Arg Arg Tyr Arg Phe
 165 170 175
 His Pro Ile Asp Pro Ser Phe Pro Glu Pro Tyr Ser Asn Trp Pro Ala
 180 185 190
 Thr Leu Arg His Pro Thr Ser Asp Gly Ser Asp Ala Lys Asp Asn Val
 195 200 205
 Lys Asp Leu Thr Thr Thr Leu Lys Ala Asp Gln Pro Asp Ile Thr Thr
 210 215 220
 Lys Thr Tyr Asn Leu Leu Thr Arg Val His Thr Trp Pro Ala Phe Ser
 225 230 235 240
 Asn His Thr Pro Gly Asp Gly Gly Ser Ser Ser Asn Ser Leu Glu Ala
 245 250 255
 Ile His Asp His Ile His Asp Ser Val Gly Gly Gly Gly Gln Met Gly
 260 265 270
 Asp Pro Ser Val Ala Gly Phe Asp Pro Ile Phe Phe Leu His His Cys
 275 280 285
 Gln Val Asp Arg Leu Leu Ala Leu Trp Ser Ala Leu Asn Pro Gly Val
 290 295 300
 Trp Val Asn Ser Ser Ser Ser Glu Asp Gly Thr Tyr Thr Ile Pro Pro
 305 310 315 320

Asp Ser Thr Val Asp Gln Thr Thr Ala Leu Thr Pro Phe Trp Asp Thr
 325 330 335
 Gln Ser Thr Phe Trp Thr Ser Phe Gln Ser Ala Gly Val Ser Pro Ser
 340 345 350
 Gln Phe Gly Tyr Ser Tyr Pro Glu Phe Asn Gly Leu Asn Leu Gln Asp
 355 360 365
 Gln Lys Ala Val Lys Asp His Ile Ala Glu Val Val Asn Glu Leu Tyr
 370 375 380
 Gly His Arg Met Arg Lys Thr Phe Pro Phe Pro Gln Leu Gln Ala Val
 385 390 395 400
 Ser Val Ala Lys Gln Gly Asp Ala Val Thr Pro Ser Val Ala Thr Asp
 405 410 415
 Ser Val Ser Ser Ser Thr Thr Pro Ala Glu Asn Pro Ala Ser Arg Glu
 420 425 430
 Asp Ala Ser Asp Lys Asp Thr Glu Pro Thr Leu Asn Val Glu Val Ala
 435 440 445
 Ala Pro Gly Ala His Leu Thr Ser Thr Lys Tyr Trp Asp Trp Thr Ala
 450 455 460
 Arg Ile His Val Lys Lys Tyr Glu Val Gly Gly Ser Phe Ser Val Leu
 465 470 475 480
 Leu Phe Leu Gly Ala Ile Pro Glu Asn Pro Ala Asp Trp Arg Thr Ser
 485 490 495
 Pro Asn Tyr Val Gly Gly His His Ala Phe Val Asn Ser Ser Pro Gln
 500 505 510
 Arg Cys Ala Asn Cys Arg Gly Gln Gly Asp Leu Val Ile Glu Gly Phe
 515 520 525
 Val His Leu Asn Glu Ala Ile Ala Arg His Ala His Leu Asp Ser Phe
 530 535 540
 Asp Pro Thr Val Val Arg Pro Tyr Leu Thr Arg Glu Leu His Trp Gly
 545 550 555 560
 Val Met Lys Val Asn Gly Thr Val Val Pro Leu Gln Asp Val Pro Ser
 565 570 575
 Leu Glu Val Val Val Leu Ser Thr Pro Leu Thr Leu Pro Pro Gly Glu
 580 585 590
 Pro Phe Pro Val Pro Gly Thr Pro Val Asn His His Asp Ile Thr His
 595 600 605
 Gly Arg Pro Gly Gly Ser His His Thr His
 610 615

<210> 17
<211> 3037
<212> DNA
<213> Halocyphina villosa

<220>
<221> CDS
<222> (260) .. (388)

<220>
<221> CDS
<222> (444) .. (494)

<220>
<221> CDS
<222> (559) .. (663)

<220>
<221> CDS
<222> (713) .. (727)

<220>
<221> CDS
<222> (779) .. (802)

<220>
<221> CDS
<222> (850) .. (975)

<220>
<221> CDS
<222> (1024) .. (1134)

<220>
<221> CDS
<222> (1188) .. (1286)

<220>
<221> CDS
<222> (1333) .. (1434)

<220>
<221> CDS
<222> (1515) .. (1575)

<220>
<221> CDS
<222> (1625) .. (1845)

<220>
<221> CDS
<222> (1973) .. (2140)

<220>
<221> CDS
<222> (2187) .. (2201)

<220>
 <221> CDS
 <222> (2249)..(2311)

<220>
 <221> CDS
 <222> (2367)..(2414)

<220>
 <221> CDS
 <222> (2463)..(2489)

<220>
 <221> CDS
 <222> (2542)..(2646)

<220>
 <221> CDS
 <222> (2697)..(2834)

<400> 17
 tggggagatg gttctatata tcaaaatgat cttctgtcct gagctttcct cgtccttggt 60
 ttcgtcttgt cagtgccgag acatgctttt attaaaccat tggcgagctg cccgcgcccc 120
 aggagatagc ataatcgcct gagaaaccta gtcgtctcat ggccgtgtaa ccgttcttgc 180
 gacttatttt cgcacttctc tcagaatata aaggcctatt gtgatacggt tcattctaacc 240
 ccagcgtccc ctccgaaag atg ggc tgc ctc tca ctc ttc gca ttc ctt act 292
 Met Gly Cys Leu Ser Leu Phe Ala Phe Leu Thr
 1 5 10
 gct tta aac tca gtt cat gcc gct gtg ggt ccc gtt acg gac tta aca 340
 Ala Leu Asn Ser Val His Ala Ala Val Gly Pro Val Thr Asp Leu Thr
 15 20 25
 ctg atc gta gat act gtc gcc ccc gac ggt gct gct ttc gcg cgg gaa 388
 Leu Ile Val Asp Thr Val Ala Pro Asp Gly Ala Ala Phe Ala Arg Glu
 30 35 40
 ggtgagactt tgcgactgta aatgccggat ttgagtttct aattataatc ttcca gcc 446
 Ala
 att gtc gtc caa gag gaa cca aac tcc gtc att ggt ccg gtc atc gta 494
 Ile Val Val Gln Glu Glu Pro Asn Ser Val Ile Gly Pro Val Ile Val
 45 50 55 60
 ggtgggtagc tacgagtctt cctccttcat ttagctcatc accaagtgat atgatattaa 554
 ttaa ggt caa aag ggg gac aac ttt cgg ctc aat gtt atc aac aat ttg 603
 Gly Gln Lys Gly Asp Asn Phe Arg Leu Asn Val Ile Asn Asn Leu
 65 70 75
 gat tct ccg aac atg cgc caa tct act tcc att cat tgg cat ggc atc 651
 Asp Ser Pro Asn Met Arg Gln Ser Thr Ser Ile His Trp His Gly Ile
 80 85 90

ttc caa gga aac ggtacgtggt atatacgata atctatctgt atccattgac Phe Gln Gly Asn 95	703
tcgaatata ggt cag aat tgg gct ggtgcgttgg ccttctgaa gcctgctcga Gly Gln Asn Trp Ala 100	757
atttatcttc ctgaattttt a gat ggc gcc gca ttc gtt aac cag Asp Gly Ala Ala Phe Val Asn Gln 105	802
gtaaggagat gttcctgcct tcgtttcccc agaactaatt atcctag tgc ccc att Cys Pro Ile 110	858
gcc ccc gga ggg gac tcg ttc ttg tac gac ttt acc gaa cct ttc cag Ala Pro Gly Gly Asp Ser Phe Leu Tyr Asp Phe Thr Glu Pro Phe Gln 115 120 125	906
act ggc aca ttt tgg tat cat tcc cat tta tca act caa tac tgc gat Thr Gly Thr Phe Trp Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp 130 135 140	954
gga ctg agg gga gca ttc gtc gttcgtttctc ttcttcatca agtcaccgct Gly Leu Arg Gly Ala Phe Val 145 150	1005
ttcttctcac ttatctag atc tac gat ccg ctc gac cct tac cgg ttg ctc Ile Tyr Asp Pro Leu Asp Pro Tyr Arg Leu Leu 155 160	1056
tac gat gtc gac gac gag tcg act gtg att act ctg gcg gac tgg tac Tyr Asp Val Asp Asp Glu Ser Thr Val Ile Thr Leu Ala Asp Trp Tyr 165 170 175	1104
cac agc tat gcg gag gac att cta atc gcg taggagattt toccaagatg His Ser Tyr Ala Glu Asp Ile Leu Ile Ala 180 185	1154
tctcctctgc ctctctgaaa tccatgaact agt gca ggc gac act atc ctc atc Ala Gly Asp Thr Ile Leu Ile 190	1208
aat ggt cac gga aga ttc gcc gga gcc ggc gga acg gca aca gaa cta Asn Gly His Gly Arg Phe Ala Gly Ala Gly Gly Thr Ala Thr Glu Leu 195 200 205 210	1256
tct gtc att act gtt gag cat gga aag cgg taggcattct ccctcggtt Ser Val Ile Thr Val Glu His Gly Lys Arg 215 220	1306
tgtagatgtg tctaatttgt gatagc tac cga ttg cga ttt gcc aat atc gct Tyr Arg Leu Arg Phe Ala Asn Ile Ala 225	1359

tgt gac cct tgg ttt gcc gtg aaa atc gat agc cat acg aac ctt cgc Cys Asp Pro Trp Phe Ala Val Lys Ile Asp Ser His Thr Asn Leu Arg 230 235 240 245	1407
ggt atc gaa gct gac ggt att act act gtg cct gtca cgggtggactc Val Ile Glu Ala Asp Gly Ile Thr Thr 250	1454
cttcaatgta ggcttaccct tagcactttc ccactctgga tcctcttatg acttcccaag	1514
atc ttt gtg ggc caa cga tat agt gtc atc ctc cat gcc aac cag cct Ile Phe Val Gly Gln Arg Tyr Ser Val Ile Leu His Ala Asn Gln Pro 255 260 265 270	1562
ggt gga aac tac t gtaagctgcc taaatgttgc atgactgtcc atgattctaa Val Gly Asn Tyr	1615
ccccgccag gg att cgg gcc gct ccg aac ggc gtg agc aat ttc gcg ggt Trp Ile Arg Ala Ala Pro Asn Gly Val Ser Asn Phe Ala Gly 275 280 285	1665
ggg atc gac tcg gct att ctc cgt tat gtt ggc gcc cca gaa gaa gag Gly Ile Asp Ser Ala Ile Leu Arg Tyr Val Gly Ala Pro Glu Glu Glu 290 295 300	1713
ccc aac act agt gag gat act cca tcc gac aca ctt caa gag cag gat Pro Asn Thr Ser Glu Asp Thr Pro Ser Asp Thr Leu Gln Glu Gln Asp 305 310 315 320	1761
ctt cac ccg ctg atc cta ccc ggc gcg cca ggc atc cac tcc cgt ggg Leu His Pro Leu Ile Leu Pro Gly Ala Pro Gly Ile His Ser Arg Gly 325 330 335	1809
gcc gcc gac gtt gtc cac acc gta tca atg gag ttt gtgagtgtgg Ala Ala Asp Val Val His Thr Val Ser Met Glu Phe 340 345	1855
cgacttttct ggcccccttt attaatataa tctgggttagg atggcgcaaa cttccaattc	1915
ctcctggatg gcgtggcctt ccagccgtgc gtcatctctt tcaaagaatt tatctag	1972
ctg acg att ttg aaa tgt agc ccg acc atg ccc gtc ctt ctg caa ata Leu Thr Ile Leu Lys Cys Ser Pro Thr Met Pro Val Leu Leu Gln Ile 350 355 360	2020
tta tcg gga gcg cag act gct aat acc ctt ctc ccg gcg gga tcc ttt Leu Ser Gly Ala Gln Thr Ala Asn Thr Leu Leu Pro Ala Gly Ser Phe 365 370 375 380	2068
atc caa gcg tcg cac aat gac atc gtg gag ctc aat ttc cca gct gtc Ile Gln Ala Ser His Asn Asp Ile Val Glu Leu Asn Phe Pro Ala Val 385 390 395	2116
aac gta gcc gct gtc ggt gga ccg tgcgtcccat ctttccttgc cagcttgaaa Asn Val Ala Ala Val Gly Gly Pro 400	2170

tttacgctct tttaga cat cca atc cat ctg tgagcgcagc gggacctttg	2221
His Pro Ile His Leu	
405	
gcttatggca tatgacttat tattagc cat ggc cat gca ttc gac gtt ata cgc	2275
His Gly His Ala Phe Asp Val Ile Arg	
410 415	
tct gct gga acg aac tcc gat aac tgg ttc aat ccg gtattttcat	2321
Ser Ala Gly Thr Asn Ser Asp Asn Trp Phe Asn Pro	
420 425 430	
tcgacttcca taagatgacg atggctcact atgggttttta ccag cct cgc aga gat	2378
Pro Arg Arg Asp	
gtc gta tcc acc ggt acc gat cct aat gac aat gtg tacgtgttc	2424
Val Val Ser Thr Gly Thr Asp Pro Asn Asp Asn Val	
435 440 445	
gctattgatt gtccgttttg atttgactgt tggacagc acc att cgc ttc cgg gcc	2480
Thr Ile Arg Phe Arg Ala	
450	
gac aac ccg tacgtaaaact gctgaatctc tcgttgtctt tggttctcat	2529
Asp Asn Pro	
455	
aatctcatca ga ggt cca tgg ttc ctt cac tgc cac att gac tgg cac ctt	2580
Gly Pro Trp Phe Leu His Cys His Ile Asp Trp His Leu	
460 465	
gaa ctc ggc ttt gct ttg gtg att gca gaa gcg cct agc gaa tgg gac	2628
Glu Leu Gly Phe Ala Leu Val Ile Ala Glu Ala Pro Ser Glu Trp Asp	
470 475 480	
agc gac att aac cct cct ggtgcgtgc ctgtgaacct tttctoccta	2676
Ser Asp Ile Asn Pro Pro	
485 490	
cacttgctaa gatcgtctta gct gcg tgg gat gac cta tgc cct acg ttc gct	2729
Ala Ala Trp Asp Asp Leu Cys Pro Thr Phe Ala	
495 500	
tgg ctt ctc ttt tac tat ttc aag ttt cct cac att ctc aac ttc aca	2777
Trp Leu Leu Phe Tyr Tyr Phe Lys Phe Pro His Ile Leu Asn Phe Thr	
505 510 515	
gat atg atg ccc tgc cgc ctg agc agc agt aat cga gtt aag aac ctc	2825
Asp Met Met Pro Cys Arg Leu Ser Ser Ser Asn Arg Val Lys Asn Leu	
520 525 530	
aac gtt gac taaggaaaaa gcaaagcaga atatgaaact ctcatattatc	2874
Asn Val Asp	
535	
tttatatcga cacattcact attcaaccta cggatttttc ctgcacctg aatttcggtg	2934

ctagatcccc atccttggtg gagtaggaaa gaaatttctt gtataaaaacc catgggttct 2994
 tctaccaata tatacataac gtccgtgggg ttagttaatt cgt 3037

<210> 18
 <211> 536
 <212> PRT
 <213> Halocyphina villosa

<400> 18
 Met Gly Cys Leu Ser Leu Phe Ala Phe Leu Thr Ala Leu Asn Ser Val
 1 5 10 15
 His Ala Ala Val Gly Pro Val Thr Asp Leu Thr Leu Ile Val Asp Thr
 20 25 30
 Val Ala Pro Asp Gly Ala Ala Phe Ala Arg Glu Ala Ile Val Val Gln
 35 40 45
 Glu Glu Pro Asn Ser Val Ile Gly Pro Val Ile Val Gly Gln Lys Gly
 50 55 60
 Asp Asn Phe Arg Leu Asn Val Ile Asn Asn Leu Asp Ser Pro Asn Met
 65 70 75 80
 Arg Gln Ser Thr Ser Ile His Trp His Gly Ile Phe Gln Gly Asn Gly
 85 90 95
 Gln Asn Trp Ala Asp Gly Ala Ala Phe Val Asn Gln Cys Pro Ile Ala
 100 105 110
 Pro Gly Gly Asp Ser Phe Leu Tyr Asp Phe Thr Glu Pro Phe Gln Thr
 115 120 125
 Gly Thr Phe Trp Tyr His Ser His Leu Ser Thr Gln Tyr Cys Asp Gly
 130 135 140
 Leu Arg Gly Ala Phe Val Ile Tyr Asp Pro Leu Asp Pro Tyr Arg Leu
 145 150 155 160
 Leu Tyr Asp Val Asp Asp Glu Ser Thr Val Ile Thr Leu Ala Asp Trp
 165 170 175
 Tyr His Ser Tyr Ala Glu Asp Ile Leu Ile Ala Ala Gly Asp Thr Ile
 180 185 190
 Leu Ile Asn Gly His Gly Arg Phe Ala Gly Ala Gly Gly Thr Ala Thr
 195 200 205
 Glu Leu Ser Val Ile Thr Val Glu His Gly Lys Arg Tyr Arg Leu Arg
 210 215 220
 Phe Ala Asn Ile Ala Cys Asp Pro Trp Phe Ala Val Lys Ile Asp Ser
 225 230 235 240
 His Thr Asn Leu Arg Val Ile Glu Ala Asp Gly Ile Thr Thr Ile Phe
 245 250 255

Val Gly Gln Arg Tyr Ser Val Ile Leu His Ala Asn Gln Pro Val Gly
 260 265 270
 Asn Tyr Trp Ile Arg Ala Ala Pro Asn Gly Val Ser Asn Phe Ala Gly
 275 280 285
 Gly Ile Asp Ser Ala Ile Leu Arg Tyr Val Gly Ala Pro Glu Glu Glu
 290 295 300
 Pro Asn Thr Ser Glu Asp Thr Pro Ser Asp Thr Leu Gln Glu Gln Asp
 305 310 315 320
 Leu His Pro Leu Ile Leu Pro Gly Ala Pro Gly Ile His Ser Arg Gly
 325 330 335
 Ala Ala Asp Val Val His Thr Val Ser Met Glu Phe Leu Thr Ile Leu
 340 345 350
 Lys Cys Ser Pro Thr Met Pro Val Leu Leu Gln Ile Leu Ser Gly Ala
 355 360 365
 Gln Thr Ala Asn Thr Leu Leu Pro Ala Gly Ser Phe Ile Gln Ala Ser
 370 375 380
 His Asn Asp Ile Val Glu Leu Asn Phe Pro Ala Val Asn Val Ala Ala
 385 390 395 400
 Val Gly Gly Pro His Pro Ile His Leu His Gly His Ala Phe Asp Val
 405 410 415
 Ile Arg Ser Ala Gly Thr Asn Ser Asp Asn Trp Phe Asn Pro Pro Arg
 420 425 430
 Arg Asp Val Val Ser Thr Gly Thr Asp Pro Asn Asp Asn Val Thr Ile
 435 440 445
 Arg Phe Arg Ala Asp Asn Pro Gly Pro Trp Phe Leu His Cys His Ile
 450 455 460
 Asp Trp His Leu Glu Leu Gly Phe Ala Leu Val Ile Ala Glu Ala Pro
 465 470 475 480
 Ser Glu Trp Asp Ser Asp Ile Asn Pro Pro Ala Ala Trp Asp Asp Leu
 485 490 495
 Cys Pro Thr Phe Ala Trp Leu Leu Phe Tyr Tyr Phe Lys Phe Pro His
 500 505 510
 Ile Leu Asn Phe Thr Asp Met Met Pro Cys Arg Leu Ser Ser Ser Asn
 515 520 525
 Arg Val Lys Asn Leu Asn Val Asp
 530 535